PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Estimating population prevalence of potential airflow obstruction using different spirometric criteria: a pooled cross-sectional analysis of persons aged 40-95 years in England and Wales
AUTHORS	Scholes, Shaun; Moody, Alison; Mindell, Jenny

VERSION 1 - REVIEW

REVIEWER	Avinesh Pillai Department of Statistics University of Auckland New Zealand
REVIEW RETURNED	06-Jun-2014

GENERAL COMMENTS	The authors have clearly stated the analysis methods used, but statistical references should be included for the tests used (Chi-square tests, analysis of variance (ANOVA), multinomial regression, and logistic regression. The results are presented clearly, given the multitude of analyses undertaken. It is reassuring to note that the datasets and statistical code used are available (from the UK Data Service and the corresponding author respectively), especially as three different statistical packages were used for the analyses. The strength and limitations are clearly stated. One minor point to note is that the authors say on page 16 line 50, 'Combining two datasets ensured a sufficient sample size to estimate prevalence, and infer valid statistical associations'. I query this statement because including the UKHLS dataset only, with 5936 participants would also ensure a sufficient sample size. However, this is a minor point. It s also good to see the authors provide context, by comparing their results with the NHANES survey for similar age groups.

REVIEWER	Hisamitsu Omori, MD. PhD Professor, Department of Biomedical Laboratory Sciences, Faculty of Life Sciences, Kumamoto University, Japan
REVIEW RETURNED	10-Jun-2014

GENERAL COMMENTS	This study demonstrates the impact of different definitions on the prevalence of potential airflow obstruction and its associations with key risk factors and comorbidities. I think this study will give us important massages in this field and have sufficiently high priority for publication in the BMJ open.

REVIEWER	Rachel Jordan University of Birmingham, UK
REVIEW RETURNED	12-Jun-2014

GENERAL COMMENTS	Although the authors have gone to much effort, this paper unfortunately does not appear to add anything novel to the field, and for this reason I do not think it should be published in its current format. The aim of the paper is to estimate the population prevalence of airflow obstruction/COPD using different criteria, using data from the latest HSE and UK household surveys. The main conclusions are that using the fixed ratio and LLN criteria provides a different picture of the size and distribution of the disease burden, finding that the fixed ratio results in higher prevalences, particularly among older people, and amongst males. All of these points have been demonstrated many times (including in England) in the papers that the authors cite. The authors also examine the association between a limited range of risk factors and airflow obstruction (with each of the definitions), but none of these results are particularly novel (older, current smoking, greater pack-year history, employed in routine occupations increased risk). In addition, the paper is very long and detailed, and has many tables with multiple comparisons in each. This makes the key messages very difficult to extract and follow. The authors should choose which are the most sensible data to show to make the points. In terms of scientific validity, the following points are also important: 1.The authors do not demonstrate that it is appropriate (or necessary) to combine the two datasets – I would like to see basic characteristics of the two datasets and possibly also assurance that the results behave similarly in both datasets. Plus also a clear table (rather than long text) of how the variables were measured in both 2.In the UKHLS patients were not asked if they had "COPD" which these days could cause problems of underestimation in the self-reported outcome measure 3.I did not follow all the statistical analysis section. There was too much detail on the basic descriptive statistics, and not enough clarity on the more complex parts. It seems that the authors were using an outcome with multiple levels – is this necessary? Does it help? It is quite confusing. 4.In the LLN model the stages were calculated in a relatively novel way. Thus introducing a further level of complexity and a different research question 5.To estimate the association with underdiagnosis (another separate research question), four outcome categories were created. First, this analysis should be looked at by a statistician as it is not as straightforward as with a binary outcome. Second, the "underdiagnosed" assumes that anyone with airflow obstruction but no diagnosis is "undiagnosed" COPD. This may not be the case – clinical symptoms should really be required – our knowledge of what it means to have airflow obstruction and no clinical symptoms is uncertain – better natural history studies are needed to find out whether they eventually progress to a clinical disease. Also, it seems confusing to create 4 categories - none, diagnosed but not obstructed, obstructed but not diagnosed, both. 6.Note that the latest NICE criteria are actually not using the predicted FEV1 <80%, but simply the fixed cut-off. 7.The use of the words "restricted" in several places to mean the

	"tightest" definition is confusing because there is also such a thing as restricted disease, which is different to obstructive disease. 8.In the results section (page 12) please make clear that you mean chronic cough and phlegm as the respiratory symptoms – it is too vague as it stands. It also seems that many diagnosed COPD patients have no symptoms/dyspnoea – is that correct? 9.Tables – in general – far too many ideas/comparisons in each table to understand easily, plus multiple models and multiple testing 10.I don't find the calculation of sensitivity and specificity necessary or helpful in this circumstance

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

The authors have clearly stated the analysis methods used, but statistical references should be included for the tests used (Chi-square tests, analysis of variance (ANOVA), multinomial regression, and logistic regression.

OUR RESPONSE: We have inserted three references in the statistical analyses section (references 40 and 42-43 in the revised manuscript). Together these references cover the four statistical tests used.

The results are presented clearly, given the multitude of analyses undertaken.
It is reassuring to note that the datasets and statistical code used are available (from the UK Data Service and the corresponding author respectively), especially as three different statistical packages were used for the analyses. The strength and limitations are clearly stated.

One minor point to note is that the authors say on page 16 line 50, 'Combining two datasets ensured a sufficient sample size to estimate prevalence, and infer valid statistical associations'. I query this statement because including the UKHLS dataset only, with 5936 participants would also ensure a sufficient sample size. However, this is a minor point.

OUR RESPONSE: We have amended the sentence (pp.16-17) as follows:

"Combining the HSE and UKHLS datasets increased statistical precision for spirometry-based estimates, particularly for population subgroups, and allowed detailed analyses to be conducted".

It's also good to see the authors provide context, by comparing their results with the NHANES survey for similar age groups.

Thank you

Reviewer: 2

This study demonstrates the impact of different definitions on the prevalence of potential airflow obstruction and its associations with key risk factors and comorbidities. I think this study will give us important massages in this field and have sufficiently high priority for publication in the BMJ open.

OUR RESPONSE: Thank you.

Reviewer: 3

Although the authors have gone to much effort, this paper unfortunately does not appear to add anything novel to the field, and for this reason I do not think it should be published in its current format. The aim of the paper is to estimate the population prevalence of airflow obstruction/COPD using different criteria, using data from the latest HSE and UK household surveys. The main conclusions are that using the fixed ratio and LLN criteria provides a different picture of the size and distribution of the disease burden, finding that the fixed ratio results in higher prevalences, particularly among older people, and amongst males. All of these points have been demonstrated many times (including in England) in the papers that the authors cite.

The authors also examine the association between a limited range of risk factors and airflow obstruction (with each of the definitions), but none of these results are particularly novel (older, current smoking, greater pack-year history, employed in routine occupations increased risk). In addition, the paper is very long and detailed, and has many tables with multiple comparisons in each. This makes the key messages very difficult to extract and follow. The authors should choose which are the most sensible data to show to make the points.

OUR RESPONSE: Thank you for these points. We respond to each in turn below.

Value of the study

The aims of the paper are as stated above. We submitted our paper to BMJ Open on the basis of the research being conducted properly, that it is scientifically credible, and reported according to appropriate guidelines. Although not novel we feel that our study adds to the COPD epidemiological research database by providing nationally representative baseline data for monitoring purposes in the UK and for facilitating comparison with international studies. It is the first time that the different definitions have been compared in nationally-representative samples, using quality-controlled spirometry data. Previous research based on English survey data (references 37, 39, 49 and 61) used Health Survey for England (HSE) 2001 data. Spirometry for both HSE 2010 and UKHLS differed substantially in the measurement protocols, equipment, and reference spirometry data from those used in HSE 2001 (and earlier years), because of major technological advances that improved quality control. More details on these are available in the HSE 2010 report:

Mindell J, Chaudhury M, Aresu M, Jarvis D. Lung function in adults. Chapter 3 in Craig R, Mindell J (eds). Health Survey for England 2010. Health and Social Care Information Centre, Leeds, 2011. (http://www.hscic.gov.uk/catalogue/PUB03023/heal-surv-eng-2010-resp-heal-ch3-func-adul.pdf)

The 2010 National Institute for Health and Care Excellence (NICE) guidance identified that 'definitive spirometry reference values were not currently available for all ethnic populations and that the European Respiratory Society (ERS) 1993 reference values are commonly used but it is recognised that these values may lead to under-diagnosis in older people and are not applicable in black and Asian populations' (NICE guidance 2010, p.74). In our study, use of the GLI 2012 equations to derive predicted values allowed us to include both an extended age-range and participants from non-white minority ethnic groups.

Risk factors

We limited the set of risk factors to those most commonly associated with airflow obstruction/COPD: age; sex; socioeconomic status; and smoking history. Variables such as limiting long-standing illness, doctor-diagnosed cardiovascular disease, and Body Mass Index were not included in regression modelling due to an inability to infer temporal associations with cross-sectional data. But we have presented these bivariate associations as Supplementary data to show higher prevalence of unfavourable health outcomes in participants within the tightest categories of FT- and LLN-defined airflow obstruction. Rather than identify a full set of risk factors associated with airflow obstruction, our aim, as in other studies, was to examine whether the associations between the four risk factors listed above and airflow obstruction varied across the different definitions. As in the study by Jordan et al (2012), we found that age- and sex-differences in risk were more marked for FT- than LLN-defined obstruction.

Length and detail

The paper is detailed as we combined two surveys to compare different definitions of airflow obstruction, and examined the sensitivity of results to including/excluding persons with diagnosed asthma (due to bronchodilators not being used before study spirometry). We defined outcomes using more than two categories to accommodate severity assessment (see Point 3 below). Both the original and revised manuscript fall below the suggested 4,000 word count: and we have made judicious use of supplementary data. We would be happy to discuss this further if the editor wishes.

In terms of scientific validity, the following points are also important:

1.The authors do not demonstrate that it is appropriate (or necessary) to combine the two datasets – I would like to see basic characteristics of the two datasets and possibly also assurance that the results behave similarly in both datasets. Plus also a clear table (rather than long text) of how the variables were measured in both.

OUR RESPONSE: Thank you. We deal with these separate points in turn below.

Combining HSE and UKHLS datasets

We have revised the manuscript to more carefully spell out the reasons for combining both datasets. A longer response is provided below, followed by the revised text.

Recent guidelines recommend adopting multidimensional assessments of airflow obstruction, taking into account symptoms such as breathlessness, exercise limitation, and exacerbations. Using the HSE 2010 we examined the associations between FT- and LLN-defined airflow obstruction and (i) breathlessness and (ii) respiratory symptoms such as chronic cough and phlegm. Our results showed the extent to which breathlessness and the presence of respiratory symptoms increased with disease-staging in consistent ways for both definitions. We hope this information is valuable to policy-makers and the research community.

Information on breathlessness and respiratory symptoms was not available for the UKHLS, the larger of the two surveys. Nevertheless, as both surveys collected lung function data using identical equipment and protocols, and we were assured of similar characteristics in both surveys (see Table 1 below), we were able to increase statistical precision for spirometry-based estimates, particularly for population subgroups, by combining both surveys.

Our revised text in the introduction (p.5) is as follows:

“Two nationally-representative samples, Wave 2 (2010-2012) of the UK Household Longitudinal Survey (UKHLS, ‘Understanding Society’) and the Health Survey for England (HSE) 2010, collected lung function data using identical measurement protocols and specialist equipment, providing an opportunity to increase statistical precision by combining both datasets”.

Survey-specific results

Providing survey-specific results would have added length and detail to the paper. We found reassuring similarities in the basic characteristics of both datasets prior to combining them. Table 1 at the end of this document compares the analytical sample across the two surveys by sex; age; smoking history; socioeconomic status; and objective measurements of lung function. We now present this as Table S1 in the revised Supplementary data. We comment on this in the first paragraph of the revised results section as follows:

"Differences between the UKHLS and HSE in terms of sex ratio, age, smoking history, NS-SEC, and objective measurements of lung function were not materially important (see online supplementary Table S1)".

We also included a survey-specific indicator in the multivariate modelling to examine potential differences in relative risk ratios (RRRs) after adjustment for the other covariates. This indicator was not statistically significant for spirometer-based outcomes (Tables 3 and 4). The two surveys were different, however, with regards to self-reported doctor-diagnosed chronic bronchitis, emphysema or COPD. We discuss this further below in response to Point 2.

Measurement of variables

We have shortened the text describing how the variables were measured whilst conforming to STROBE guidelines. The edits are shown as tracked changes in the revised manuscript.

2.In the UKHLS patients were not asked if they had "COPD" which these days could cause problems of underestimation in the self-reported outcome measure

OUR RESPONSE: Thank you. We have added this point to the strengths and limitations section of the discussion. The revised text (pp.17-18) is as follows:

"The list of health conditions in the UKHLS interview programme included chronic bronchitis and emphysema but not COPD, leading to potential underestimation of self-reported physician-diagnosed COPD".

3.I did not follow all the statistical analysis section. There was too much detail on the basic descriptive statistics, and not enough clarity on the more complex parts. It seems that the authors were using an outcome with multiple levels – is this necessary? Does it help? It is quite confusing.

OUR RESPONSE: We have expanded the statistical analyses section to provide more clarity on multinomial logistic regression. The revised text (page 10) is as follows:

"Multinomial logistic regression generalises logistic regression to outcomes with more than two possible discrete outcomes. The RRR is interpreted as the relative risk of one outcome in relation to the reference category for a specified category of an independent variable compared with the reference".

We have also added an explanatory footnote to Tables 3 and 4:

"The RRR is interpreted as the relative risk of one outcome in relation to the reference category for a specified category of an independent variable compared with the reference category for that independent variable. Using FT stage I as an example, the RRR for males vs. females is interpreted as the relative risk for FT stage I vs. non-obstruction for males compared with the analogous relative risk for females, adjusted for the other variables in the model".

Outcomes with multiple levels

FT- and LLN-outcomes were defined using multiple levels to examine how the associations with risk factors / comorbidities changed as the definition of obstruction became 'tighter'. Our results showed that variables such as breathlessness, the presence of respiratory symptoms, and self-reported cardiovascular disease increase with disease-staging.

4.In the LLN model the stages were calculated in a relatively novel way. Thus introducing a further level of complexity and a different research question.

OUR RESPONSE: As described in the paper by Swanney et al:

Swanney MP, Ruppel G, Enright PL, Pedersen OF, Crapo RO, Miller MR et al. Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction. Thorax 2008; 63(12):1046-1051.

Guidelines from the National Lung Health Education Programme in the USA (NLHEP) require both the FEV1/FVC and FEV1 to be below the LLN. We chose to split participants with FEV1/FVC < LLN into two categories: (FEV1 ≥ LLN and FEV1 < LLN) to examine possible heterogeneity among participants with FEV1/FVC < LLN. Our revised text on page 7 is as follows:

"To examine possible heterogeneity among participants with FEV1/FVC < LLN, disease stage was defined by FEV1 relative to LLN as follows: stage I (FEV1/FVC

Ferguson GT, Enright PL, Buist AS, et al. Office spirometry for lung health assessment in adults: A consensus statement from the National Lung Health Education Program. Chest 2000; 117(4): 1146-61.

5.To estimate the association with underdiagnosis (another separate research question), four outcome categories were created. First, this analysis should be looked at by a statistician as it is not as straightforward as with a binary outcome. Second, the "underdiagnosed" assumes that anyone with airflow obstruction but no diagnosis is "undiagnosed" COPD. This may not be the case – clinical symptoms should really be required – our knowledge of what it means to have airflow obstruction and no clinical symptoms is uncertain – better natural history studies are needed to find out whether they eventually progress to a clinical disease. Also, it seems confusing to create 4 categories - none, diagnosed but not obstructed, obstructed but not diagnosed, both.

OUR RESPONSE: Thank you for raising these points. We respond to each below.

Multinomial logistic regression

The corresponding author is an experienced social statistician, and is comfortable with multinomial logistic regression. This statistical technique avoids the possible oversimplifications of using an outcome with only two categories, and so can accommodate outcome variables capturing severity assessment (e.g., none, mild, moderate, severe). Similar applications of multinomial logistic regression include the following papers:

Cable N, Kelly Y, Bartley M, et al (2014). Critical role of smoking and household dampness during childhood for adult phlegm and cough: a research example from a prospective cohort study in Great

Britain. BMJ Open 2014;4:e004807 doi:10.1136/bmjopen-2014-004807.

Paige E, Korda RJ, Banks E, et al (2014). How weight change is modelled in population studies can affect research findings: empirical results from a large-scale cohort study. BMJ Open 2014;4:e004860 doi:10.1136/bmjopen-2014-004860.

We have also added a web reference showing how users can interpret Stata output from a multinomial logit model (see response to Reviewer 1) and have expanded the statistical analyses section to further clarify this technique (see response to Reviewer 3, Point 3).

Four categories combining doctor-diagnosed COPD and objective spirometry

To examine under-diagnosis we created a four-category outcome variable combining reported diagnosed COPD and spirometric criteria as follows:

- Neither diagnosed nor spirometrically-defined obstruction
- Physician-diagnosed COPD but no obstructive spirometry
- Spirometrically-defined but no diagnosed COPD
- Both diagnosed and obstructive spirometry.

We feel that this four-fold classification is a sensible way to cross-classify self-reported doctor-diagnosed COPD and objective spirometry; with the third category shedding light on potential under-diagnosis.

Under-diagnosis in COPD

We agree with Reviewer 3 that the category ‘spirometrically-defined but no diagnosed COPD’ does not necessarily indicate underdiagnosis. We have added this important point to the revised discussion (page 19) as follows:

“Spirometrically-defined airflow obstruction but no diagnosed COPD does not necessarily indicate under-diagnosis. Definitive diagnosis requires further information on all relevant clinical factors, particularly respiratory symptoms and smoking history, as well as post-bronchodilator spirometry.”

6.Note that the latest NICE criteria are actually not using the predicted FEV1 <80%, but simply the fixed cut-off.

OUR RESPONSE: Thank you for raising this point. We have revised the text throughout to clearly distinguish between the 2004 guidance – which required predicted FEV1<80% (but did not specify whether spirometry measurements should be made pre- or post-bronchodilator) – and the 2010 guidance which uses a post-bronchodilator FEV1/FVC < 0.70.

7.The use of the words “restricted” in several places to mean the “tightest” definition is confusing because there is also such a thing as restricted disease, which is different to obstructive disease.

OUR RESPONSE: Thank you for this point. We have changed the text from restricted to “tightest” definitions.

8.In the results section (page 12) please make clear that you mean chronic cough and phlegm as the respiratory symptoms – it is too vague as it stands. It also seems that many diagnosed COPD patients have no symptoms/dyspnoea – is that correct?

OUR RESPONSE: Thank you. In the methods section we defined the presence of respiratory symptoms as “usually coughing first thing in the morning, for at least 3 months a year, and bringing up phlegm from the chest most days for 3 consecutive months in a year.” For clarity, we now repeat this in the results section as follows:

“The prevalence of respiratory symptoms (chronic cough and phlegm) was 13.7%, 10.2%, and 11.3% among participants classed as having airflow obstruction according to diagnosed COPD, FT, and LLN respectively“.

Among participants with doctor-diagnosed COPD, 13.7% reported the presence of respiratory symptoms (chronic cough and phlegm) and 34.8% had a score of 3+ on the MRC breathlessness scale. These estimates are shown in Table S3 (Supplementary data) and were presented at the end of the 2nd paragraph in the results section.

9.Tables – in general – far too many ideas/comparisons in each table to understand easily, plus multiple models and multiple testing

OUR RESPONSE: Both the original and revised manuscript fall below the suggested 4,000 word count: and we have made judicious use of supplementary data. We have set out our rationale for the analyses we have conducted. We would be happy to discuss this further with the editor if necessary.

10.I don’t find the calculation of sensitivity and specificity necessary or helpful in this circumstance

OUR RESPONSE: We calculated estimates of sensitivity and specificity to demonstrate that two definitions producing similar prevalence may not necessarily identify the same people. For example, the prevalence of FT- (stage II+) and LLN-defined airflow obstruction among participants aged 65-95 was 19.1% and 15.8% respectively. The set of estimates provided in Table 2 enable readers to examine the agreement between the definitions. False positive rates were similar (8.9% vs. 5.2%); but false negative rates were divergent (26.7% vs. 39.1%). Among participants aged 65-95, there were 167 participants defined as FT stage II+ but non-obstructed using LLN, compared with 95 participants who were LLN but non-obstructed using FT stage II+. We hope this information is valuable to policy-makers and researchers.

Table 1: Participant characteristics for the analytical sample and by survey

(HSE and UKHLS) HSE UKHLS
N 7,879 1,943 5,936
Male, n (%) 3335 (46.8) 824 (48.4) 2511 (46.2)
Age-group, n (%):
40-54 3472 (46.6) 868 (45.8) 2604 (46.9)
55-64 2072 (24.8) 497 (24.2) 1575 (25.0)
65-74 1557 (17.4) 369 (17.3) 1188 (17.5)
75-95 778 (11.1) 209 (12.6) 569 (10.6)
Mean age, years (SD) 57.6 (12.3) 57.9 (12.5) 57.5 (12.2)
Smoking status, n (%):
Current 1198 (16.6) 254 (14.5) 944 (17.3)
Ex-regular 2547 (31.7) 659 (33.1) 1888 (31.3)
Never 4134 (51.7) 1030 (52.4) 3104 (51.5)
Pack-years, n (%):
0-0.9 4299 (53.9) 1082 (55.0) 3217 (53.5)
1-19.9 1905 (24.3) 493 (25.1) 1412 (24.0)

20-49.9 1318 (17.2) 283 (15.0) 1035 (17.9)
50+ 345 (4.6) 80 (4.7) 265 (4.5)
NS-SEC:
Professional 3050 (36.5) 772 (36.1) 2278 (36.6)
Intermediate 1859 (23.4) 452 (23.6) 1407 (23.3)
Routine 2705 (36.9) 709 (39.8) 1996 (36.0)
Missing 265 (3.2) 10 (0.5) 255 (4.1)
Lung function (%-of-predicted),
Mean (SD):
FEV1 92.0 (16.5) 91.9 (16.4) 92.0 (16.5)
FVC 97.1 (15.0) 97.2 (15.0) 97.1 (15.1)
FEV1/FVC 94.2 (9.7) 94.0 (9.9) 94.3 (9.7)

Abbreviations: FEV1 = forced expiratory volume in one second; FVC, forced vital capacity; HSE = Health Survey for England; NS-SEC = National Statistics Socio-Economic Classification; SD = standard deviation; UKHLS = United Kingdom Household Longitudinal Study;

This Table has been added to the Supplementary Data (Table S1).